DOCUMENT RESUME

ED 177 470 CS 004 968

AUTHOR	Anderson, Richard C.; Freebody, Peter
TITLE	Vocabulary Knowledge and Reading. Reading Education Report No. 11.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Aug 79
CONTRACT	400-76-0116
NOTE	52p.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Intelligence Factors; *Reading Comprehension; *Reading Processes; *Reading Research; *Vocabulary; *Vocabulary Development
IDENTIFIERS	*Center for the Study of Reading IL

ABSTRACT

This report summarizes information presented in a technical report concerning the role of vocabulary knowledge in reading comprehension and is designed to be read by educators not directly involved in research themselves. It states that while an assessment of the number of meanings a reader knows enables a remarkably accurate prediction of an individual's ability to comprehend discourse, the reasons why word knowledge correlates with comprehension cannot be determined satisfactorily without improved methods of estimating the size of people's vocabularies. It suggests that improved assessment methods depend upon thoughtful answers to questions concerning what a word is, what it means to know the meaning of a word, and the most efficient way of estimating vocabulary size from an individual's performance on a sample of words. (MKM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED177470

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CENTER FOR THE STUDY OF READING

Reading Education Report No. 11

VOCABULARY KNOWLEDGE AND READING

Richard C. Anderson and Peter Freebody

University of Illinois at Urbana-Champaign

August 1979

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported in part by the National Institute of Education under Contract No. US-NIE-C-400-76-0116.

CS004968

Vocabulary Knowledge and Reading

Our aim in this paper is to summarize what is known about the role of vocabulary knowledge in reading comprehension. Though word identification skills are important in reading, this paper is concerned exclusively with knowledge of word meanings. An assessment of the number of meanings a reader knows enables a remarkably accurate prediction of this individual's ability to comprehend discourse. That this is true may seem self-evident; why it is true is poorly understood. Determining why is important because what should be done to build vocabulary knowledge depends on why it relates so strongly to reading. The deeper reasons why word knowledge correlates with comprehension cannot be determined satisfactorily without improved methods of estimating the size of people's vocabularies. Improved assessment methods hinge, in turn, on thoughtful answers to such questions as what is a word, what does it mean to know the meaning of a word, and what is the most efficient way of estimating vocabulary size from an individual's performance on a sample of words.

Vocabulary Knowledge and Linguistic Competence

Measures of vocabulary knowledge are potent predictors of a variety of indices of linguistic competence. The strong relationship between vocabulary and general verbal ability is one of the most robust findings in the history of intelligence testing. Terman (1918), for instance, reported a correlation of .91 between mental age (as assessed by the Stanford Revision of the Binet-Simon Scale) and the vocabulary subscale.

On this basis he suggested that the vocabulary measure alone constitutes a good estimate of performance on the entire scale and thus could be used as a short measure. Since then, this suggestion has been tested with various age groups. Table 1 summarizes representative evidence. In these studies, correlations between vocabulary subtest scores and total test scores on a number of different aptitude tests have ranged from .71 to .98.

Insert Table 1 about here

An equally consistent finding has been that word knowledge is strongly related to reading comprehension. Davis (1944, 1968) factor analyzed nine comprehension tests and found a main factor for word knowledge on which a vocabulary test loaded about .8. In the years that followed, many studies have shown that a major factor in reading comprehension is vocabulary knowledge (e.g., Botzum, 1951; Michael, Zimmerman, & Guilford, 1951; Wrigley, Saunders, & Newhaus, 1958; Clark, 1972). The range of factor loadings for vocabulary tests in these studies was .41 to .93. These findings indicate the need for a central role for word knowledge in any model of reading comprehension.

Analyses of readability (cf. Bormuth, 1966) also demonstrate the preeminent role of word knowledge. In a study of the factors that make prose difficult to read, Coleman (1971) examined morphological, syntactic, and semantic properties of words and sentences. While he found sentence complexity to be a fairly important variable, he was able to conclude that

"any measure of word complexity (number of letters, morphemes, or syllables; frequency of usage) will account for about 80% of the predicted variance" (p. 184). Klare (1974-1975), in a review of readability, also concluded that a two-variable formula is sufficient for most practical purposes: one variable relates to word difficulty and the other to syntactic or sentence difficulty. He went on to conclude that the word variable is consistently more highly predictive of difficulty than is the sentence variable. As would be expected, an index of vocabulary difficulty has always been given the heaviest weight in readability formulas.

Why is Vocabulary Knowledge a Major Factor in Linguistic Competence?

There are three more or less distinct views of why vocabulary knowledge is such an extraordinary correlate of linguistic competence. We will call the first the instrumentalist position: Individuals who score high on a vocabulary test are likely to know more of the words in most texts they encounter than low scoring individuals. The heart of the instrumentalist hypothesis is that knowing the words enables text comprehension. In other words, this hypothesis claims that vocabulary knowledge is directly and importantly in the causal chain resulting in text comprehension. Unlike the two positions described below, the instrumentalist hypothesis has nothing to say about where vocabulary knowledge comes from, but only that, once possessed, it helps the reader understand text.

According to the second position, vocabulary tests measure verbal aptitude. A person who scores high on such a test has a quick mind. With

the same amount of exposure to the culture, this individual has learned more word meanings. He or she also comprehends discourse more readily than the person who scores low on a vocabulary test. The essential claim of the aptitude hypothesis is that persons with large vocabularies are better at discourse comprehension because they possess superior mental agility. A large vocabulary is not conceived to be involved in a direct way in better text understanding in this model. Rather vocabulary test performance is merely another reflection of verbal ability and it is verbal ability that mainly determines whether text will be understood.

The third position is the knowledge hypothesis. Performance on a vocabulary test is seen as a reflection of the extent of exposure to the culture. The person who scores high has deeper and broader knowledge of the culture. The essential idea is that this knowledge is crucial for text understanding. Rather than being directly important, possessing a certain word meaning is only a sign that the individual may possess the knowledge needed to understand a text. For instance, the child who knows the word mast is likely to have knowledge about sailing. This knowledge enables that child to understand a text that contains sentences which do not even involve the word mast, such as, "We jibed suddenly and the boom snapped across the cockpit."

Of course, jibe, boom, and cockpit are specialized words, too. It might be wondered whether the instrumental hypothesis and the knowledge hypothesis are really different. Stong versions of the two positions

are distinguishable, at least. The instrumental position, as we choose to characterize it, stresses individual word meanings. The knowledge view emphasizes conceptual frameworks or "schemata;" individual word meanings are merely the exposed tip of the conceptual iceberg.

Instructional Implications of Different Hypotheses About Vocabulary Knowledge

Which of the three positions is most tenable? The main point to be made is that there are neither the theoretical tools nor the data to justify a conclusion at the present time. A second important point is that it would be naive, indeed, to begin with the idea that one of the positions will turn out to be entirely right and the other two entirely wrong. However, it is important to know which of the three hypotheses about vocabulary knowledge is most nearly correct because each one has radically different implications for reading instruction. At one extreme, some who endorse the verbal aptitude hypothesis are fatalistic about whether any environmental factor can have a major influence on children's reading. They tend to recommend family planning instead of curriculum innovation as the final solution to the reading problem.

Of course the verbal aptitude position does not require the belief that heredity is predominant. Alternatively, there are those who maintain that verbal ability grows in proportion to the volume of experience with language. The greater the opportunities to use language the faster and more efficient become the elemental processing operations. In turn, speed

and efficiency permit greater benefit from each successive language encounter. More detailed accounts of this sort of position can be found in the well-known paper by La Berge and Samuels (1974) and a recent paper by Perfetti and Lesgold (in press).

The latter formulation of the verbal aptitude hypothesis leads to the recommendation that educators should try to maximize the amount of reading children do. However, this is not very newsworthy. It is a practice that would be endorsed no matter what the theoretical persuasion. The distinctive emphasis in the verbal aptitude position is on speed and efficiency of processing. This emphasis gives rise to the recommendation that beginning readers and poor readers receive extensive drill and practice on the "fundamentals" of reading. According to Perfetti and Lesgold (in press), the drill activities should include even more practice than typically provided in word vocalization, more practice in speeded word recognition, and more practice in immediate memory for the literal content of text. It should be noted that these suggestions are offered in the spirit of a hypothesis. Perfetti and Lesgold acknowledge that, so far at least, attempts to facilitate text comprehension by providing speeded word drills have not proved very successful (see especially Fleisher & Jenkins, 1977).

While, like everyone else, the advocate of the instrumental hypothesis favors lots of reading and varied language experience, the distinctive feature of this view is that it invites direct vocabulary building

exercises. Becker (1977) has argued strongly for the instrumentalist position. He maintains that once decoding skills have been mastered, the chief remaining factor in determining whether a child will be a successful reader is vocabulary knowledge. He claims that schools have never had reading programs that systematically build vocabulary. Children from middle class backgrounds pick up word meanings anyway. But the same is less true, Becker argues, of children coming from lower class homes, which often fail to provide support for the continuous vocabulary and concept growth important to school work. Consistent with this assumption is some recent research of Hall and Tirre (1979), who found that lower class parents, particularly lower class Black parents, use substantially fewer of the words found in standardized intelligence tests when speaking with their children than do middle class parents.

Becker proposes a reading curriculum in which every child would learn about 7,000 basic words from direct instruction. The figure 7,000 comes from one estimate of the number of basic words known by the average high school senior (Dupuy, 1974). Becker acknowledges that there are families of words with related meanings, thereby permitting the child some generalization beyond the words that are specifically taught. By and large, though, he believes that learning one vocabulary item gives little advantage in learning the next one. For instance, he illustrates instruction on the following set of unrelated words: help, support, insist, toil, resist, recognize, assist. Even the so-called "concept side" of the instruction

entails a component analysis of isolated words. So if this assumption is correct, direct teaching of a vocabulary of even 7,000 basic words would be an enormous task. Becker estimates that about 25 basic words would have to be taught per week from the third through the twelfth grade (p. 539).

The distinctive curriculum implication of the knowledge hypothesis is that generally new vocabulary ought to be learned in the context of acquiring new knowledge (cf. Goodman, 1976, p. 487). Every serious student of reading recognizes that the significant aspect of vocabulary development is in the learning of concepts not just words. The additional point that the knowledge position brings to the fore is that concepts come in clusters that are systematically interrelated. Returning to an earlier example, the concept of mast cannot be acquired independently of concepts such as boat and sail. Thus, it would seem to be sensible for people to learn the jargon in the context of learning about sailing and the anatomy of sailboats. According to the knowledge hypothesis, if a child were really naive, trying to teach a single sailing concept and word in isolation from the set of related concepts and words would be inefficient in the best case and completely fruitless in the worst case.

The more general point about the role of knowledge in vocabulary learning is intuitively clear. Imagine you wished to teach some French vocabulary to, let us say, two groups of English-speaking Canadian children, evenly matched on aptitude and achievement. One group is from a downtown urban area; the other is from a small fishing village. The body of words you wish to teach is concerned with fishing (trawlers, rods, nets, casting,

bait, currents, etc.). Would you expect one group to learn the words more quickly and easily than the other? Why? We do not know of actual research that has dealt systematically with these questions. One somewhat relevant study was carried out by Allen and Garton (1968). They found that physics students were much better than art students in recognizing physics words. They concluded that, for art students, physics words are semantically indistinct and thus have to be recognized on a more piece-meal basis. Familiarity with an area of knowledge increased the familiarity of the relevant words.

Knowledge can be sliced in various ways. Thus far in this section, we have considered sets of words related because they are used in talking about the same topic. Words may also be conceptualized in terms of families related to one another because they convey related sets of distinctions. Consider an example involving verbs of visual perception.[1] The basic verb is see. If you notice that look involves a deliberate act of seeing, it can then be appreciated that glimpse refers to a short act of seeing whereas glance refers to a short act of looking. Stare, on the other hand, refers to a prolonged act of looking. The variations in sense among these verbs can be understood in terms of just two semantic features, intention and duration. Further distinctions would be required to encompass other verbs of visual perception such as notice and examine.

We would consider that a lesson that helped children sharpen and extend the distinctions involved in visual perception words to be consistent

with the spirit of the knowledge position. What the knowledge position would not countenance is a separate vocabulary lesson that included glance, mast, and a miscellany of other words. Herein lies a difference from the instrumentalist position, which does not seem to us to preclude exercises involving lists of unrelated words.

Johnson and Pearson's (1978) book, Teaching Reading Vocabulary, appears to represent predominantly the knowledge position, though it is an eclectic treatment that also reflects influences from the other two views. Johnson and Pearson advocate teaching a basic sight vocabulary using "intensive direct instruction in the early grades and with older children who do not read well" (p. 28). They also endorse both direct and indirect means for teaching phonics, promoting morphological analysis, causing vocabulary knowledge to expand, and teaching the use of the dictionary and thesaurus. Johnson and Pearson devote a chapter to the use of contextual clues to figure out the meanings of unfamiliar and ambiguous words. Otherwise most of the exercises and games suggested throughout the book involve groups of words outside the context of stories or text book chapters. However, the words usually involve sets of interrelated distinctions such as were illustrated above with verbs of visual perception. Almost every activity was designed to expand children's sensitivity to these distinctions. There is an apparent discrepancy between the goals of the activities, which are concerned with conceptual distinctions and relations, and the format of the activities,

which is based largely on isolated words. If the knowledge perspective were strictly adhered to, vocabulary instruction would not be thought of as a separate subject in school.

For the sake of clarity of exposition, we have presented the aptitude, instrumental, and knowledge positions in uncomplicated and somewhat overdrawn form. We must emphasize again that no serious scholar in reading or related fields rigidly adheres to any one of these positions. In particular Hunt (1978), who is identified with the aptitude hypothesis, has explicitly and emphatically stated that vocabulary size also is a reflection of an individual's accumulated knowledge of the world. Becker, whom we labeled an instrumentalist, heartily endorses some of the implications of both the aptitude and the knowledge views. Reading has been a fractious field. If a policy were followed of avoiding controversy where none genuinely exists, the quality of intellectual exchange and the sociopolitical climate might improve to the point where someone within the next decade could write a book entitled "Learning to Read: The Great Consensus."

What Does It Mean to Know the Meaning of a Word?

It is not clear that, if Ludwig Wittgenstein and Bertrand Russell were left alone in a room for three hours, they could decide that they really knew the meaning of dog. As Labov (1973) said, "words have often been called slippery customers, and many scholars have been

distressed by their tendency to shift their meanings and slide out from under any simple definition" (p. 341).

An ordinary adult engaging in an ordinary conversation will be absolutely sure he knows the meaning of almost all of the words he hears. Notice that the restriction to ordinary use is an important aspect of this confidence. Consider the term gold, for example. The person who is sure he knows the meaning of this word in an ordinary use will quickly retreat when in the company of jewelers, mining engineers, geological survey assayists, or metalurgists.

What does a person know when he knows the meaning of a word in its ordinary, every-day, garden-variety sense? This issue is addressed in what we will refer to as the Standard Theory of semantics, according to which the meaning of a word can be analyzed into features (also called components, attributes, or properties), each of which represents one of the distinctions conveyed by the word. Necessary or essential features are usually distinguished from features that are merely characteristic. For instance, having a back could be said to be a necessary feature of chair since an object that is otherwise a chair except for the lack of a back is really a stool instead of a chair. On the other hand, the ability to fly is only a characteristic feature of bird since some birds (penguins) don't fly at all and others (chickens) do so very poorly.

To define a term, in the strong sense, is to list the features necessary to capture the essence of the thing (or event or quality) designated by the word. Saying this another way, a proper definition

indicates the attributes a thing must have in order to be designated by a word; if any of these necessary properties were missing that word would not apply. Before we choose this as our criterion in the testing of children's word knowledge, however, we might wish to examine how well it applies to adults' normal use and understanding of words.

How able are people to define the words they are sure they know? "Not very" is the answer if one insists upon the strong sense of define. Consider gold again. Upon being asked to define gold, the ordinary citizen might say that gold (a) is precious, (b) is a metal, and (c) that it has a particular yellowish (i.e., golden) hue. The problem is that none of these is a necessary feature. Not all gold is a golden color. If, say, the Chinese were to discover a mountain of gold, the substance would no longer be precious. Not even the attribute of being a metal can be considered to be an eternal, immutable property of gold for, unlikely though it is, there might be a scientific breakthrough in which it was discovered that gold is not a metal. If either of these things were to happen one would be forced to conclude that gold did not exist. As Putnam (1975) has noted, this is a very odd conclusion, because there would still be this "stuff" lying around that people used to call gold. We have a right to be suspicious of a semantic theory that backs us into such a peculiar corner.

There are other serious problems with Standard Theory. Notably, the members of a class called by the same name frequently do not all

share a single set of common properties. Wittgenstein (1953; see also Rosch, 1973; Rosch & Mervis, 1975) argued that things designated by the same word generally are related by "family resemblance." He intended an analogy to a human family whose members look and act alike. Mother and one son may have a prominent nose. Father and daughter may have the same hair color. And so on. But there may be no single respect in which they are all alike, no single feature which they all share. Wittgenstein claimed family resemblance was the most accurate characterization of the relationships among the various uses of most common words. To illustrate his point, he analyzed uses of the term game, noting the similarities and differences between team games, board games, and children's games.

A great deal more could be said about semantic theory. (For authoritative, current treatments, see Clark & Clark, 1977, especially chapters 11-14; Fillmore, 1975; and Miller & Johnson-Laird, 1976.) The main point of this brief excursion into the meaning of meaning is to caution against holding up a standard of word comprehension for children that adults could not meet.

Depth of Word Knowledge

It is useful to distinguish between two aspects of an individual's vocabulary knowledge. The first may be called "breadth" of knowledge, by which we mean the number of words for which the person knows at least some of the significant aspects of meaning. Later sections of this paper will be concerned mainly with breadth of knowledge.

Treated in this section is a second dimension of vocabulary knowledge, namely the quality or "depth" of understanding. We shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances.

Eve Clark (1973) has marshalled an array of evidence which shows that the meaning a young child has for a word is likely to be more global, less differentiated than that of an older person. With increasing age, the child makes more and more of the adult distinctions. In other words, when first acquired, the concept a child has for a word need not include all of the features of the adult concept. Eventually, in the normal course of affairs, the missing features will be learned. Most of the research done to date points to the conclusion that there is progressive differentiation of word meanings with increasing age and experience.

Just one illustration will be provided of the kind of evidence that supports this conclusion. Gentner (1975) completed a theoretical analysis of verbs of possession which indicated that buy, sell, and spend entail a more complex set of distinctions than give and take. Notice that giving involves the transfer of something from one person to another. Selling likewise involves the transfer of something from one person to another but it involves an additional transaction as well, the transfer of money from the buyer to the seller. The complimentary relationship holds between buying and taking.

Gentner expected children to acquire the full, adult meanings of these verbs in order of complexity. Children ranging from four to eight years of age were asked to make dolls act out transactions from directions involving each verb. For example, the children were requested to "make Ernie sell Bert a (toy) car." The four-year-olds performed flawlessly with directions containing give and take, but never correctly executed instructions that involved spend, buy, or sell. The eight-year-olds exhibited nearly perfect understanding of every direction except the ones containing sell. Overall, the results were exactly as expected: The adult meanings of verbs of possession are acquired in order of complexity.

There is a substantial body of literature on selected vocabulary of children from about two through eight years of age. The literature involving older children and adults is meager. In our judgment, people's vocabulary knowledge continues to deepen throughout their lifetimes; that is, that as they grow older, most people continue to learn nuances and subtle distinctions conveyed by words that in some sense they have known since childhood. There is no hard data to support this conjecture. However, an illustration will show that many adults still have something to learn about even fairly common words. It is easy to find educated adults who confuse infer and imply. A person will say something along the lines, "I intended, by stating these arguments, to infer that . . ." Of course, this individual should have said imply. Speakers imply: Listeners infer. The complication, which no doubt makes the distinction difficult, is that

speakers may report inferences they have made as well as get implications across to listeners.

Breadth of Word Knowledge

We have characterized breadth of knowledge as the number of words for which a person knows at least part of the meaning. It is disturbing to examine available estimates of the average vocabulary size of various age groups. Table 2 summarizes studies that have been carried out to estimate total basic or "root" word knowledge. It can be seen that the estimates vary wildly.

Insert Table 2 about here

There are important practical reasons for attempting to make accurate assessments of total word knowledge. Language and reading programs aim to increase students' vocabularies. The number of words presented to students varies, in part, according to what is regarded as the most authoritative thinking and research on vocabulary size and growth (Clifford, 1978). More reliable estimates would indicate the appropriateness of the assumptions of a program, and perhaps highlight periods of growth to be capitalized upon. More generally, reliable estimates would indicate whether direct language instruction can plausibly account for a substantial proportion of the child's language growth, or whether word knowledge is acquired for the most part independently of formal instruction. To refer again to a concrete proposal, Becker's (1977) idea that underachieving

children should be taught via direct instruction the vocabulary most high school seniors possess would be difficult, but perhaps feasible, if the children had to learn 25 new words a week. It would be out of the question if they had to learn 25 words each school day. Or would it?

Recently, the distinguished psychologist, George Miller (1978) stated:

> Although the rapid rate of syntactic acquisition has inspired much respectful discussion in recent years, the rate of lexical growth is no less impressive. The best figures available indicate that children of average intelligence learn new words at a rate of more than 20 per day. It seems necessary to assume therefore, that at any particular time they have hundreds of words roughly categorized as to semantic or topical relevance but not yet worked out as to precise meaning or use. (p. 1003)

Miller did not specify whether or not he was referring to basic words. If he was, then he is positing a mean annual word acquisition rate of over seven thousand words, or about fifty thousand over the elementary and middle school years. This seems unlikely even in the light of the highest estimates summarized in Table 2. He may have been including compounds and derivatives; however, to our knowledge, no systematic examination of children's ability to understand these forms has been completed. We are puzzled to know what are the "best figures available."

In 1940, Seashore and Eckerson remarked that, even though the field of vocabulary testing is a "fairly old one" (p. 35), substantial problems of measurement remained. By now, in the time span of educational research, we might want to call the field "ancient," and virtually all of those original problems persist. We will present some of the central issues in broad-gauged measurement of word knowledge. The discussion of these issues will be concerned largely with breadth of knowledge and will reveal many of the reasons why estimates of vocabulary size have fluctuated so much.

Selecting a Sample of Words

There are many thousands of words in English. Dupuy (1974), the author of a recent, thorough study of word knowledge, estimates that there are about a quarter of a million main entries in Webster's Dictionary (1961). Clearly we cannot test somebody on all the words in the dictionary, so how can we select a group of words that will yield the most reliable information? Three questions need to be asked: First, what counts as a word? Second, how should words be selected from the dictionary? Third, how large does a sample of words need to be so that the test is as reliable and informative as possible and yet still economical to administer?

In determining what is to count as a word, the researcher needs to decide whether or not it is of interest to discern the student's ability to use derivatives and compounds (plurals, participles, tense markers, comparatives, etc.). Some authors, notably Seashore (1933), have preferred

to calculate separate estimates for "special" terms and derivatives. Others, for example Dupuy (1974), have attempted to concentrate solely on "basic" words. Dupuy sampled randomly from Webster's Third New International Dictionary (1961) and then applied three criteria to each word selected: The word had to be a main entry, a single word form (i.e., not a derivative or compound), and could not be technical, slang, foreign, or archaic. By this method, Dupuy estimated that English contains 12,300 "basic" words.

The systematic nature of this sampling creates its own equally systematic biases. Some children may have acquired the generative rule, for, say, negation by prefix, such as in unable and dishonest, and others may not have (Silvestri and Silvestri, 1977). Do we wish to exclude this element of vocabulary knowledge from the measure? Adults acquire a number of special or technical terms in their areas of expertise or interest, so exclusion of technical terms denies many people the opportunity of indicating their knowledge of a large number of words.

Many researchers opt for a random sample. Once a random sample of words has been selected, a test is constructed to assess how many of the words a person knows. Then, in principle, estimating the person's vocabulary size is straightforward. For instance, Dupuy's Basic Word Vocabulary Test contains 1% of the 12,300 basic words he calculated are in Webster's. Therefore, the absolute size of the basic word vocabulary can be approximated by multiplying the score on this test by 100. A person whose score is 60, after correction for guessing, would be judged to have a basic vocabulary of 6,000 words.

One disadvantage of this method is self-evident. Estimated vocabulary size depends heavily on the size of the dictionary. With respect to Dupuy, while he sampled initially from a large unabridged dictionary, a word had to appear as a major entry in each of three other smaller dictionaries in order to be counted as a basic word. A total of 979 words, 41% of the sample, were discarded on the basis of this rule. The result was a very conservative estimate of the number of basic words in American English and is one reason Dupuy's estimates of basic vocabulary size are so much smaller than those of other investigators. Of course, many of these words were very rare, but others such as cloudlet, escaping, breezes, invited, starling, and unilateral would be familiar to most people.

Already discussed is the issue of what to do with derivative and compound forms. A liberal policy of selecting words will lead to large estimates of vocabulary size. A conservative policy will produce smaller ones. Dupuy was conservative. He eliminated 7.7% of the words in his sample on the grounds that they were compounds or derivatives, including a great many familiar ones, such as grandchild, package, and toothache. So even such a random sample might not represent the language at large.

A further problem is that projecting a vocabulary size from performance on a random sampling of words is inefficient. If the subject provides the meaning of bibulous, then using up test time by asking for the meaning of bicycle is wasteful. When estimating people's total vocabulary size is

the researcher's major aim, then efficiency of items covered per unit of examinee time is an important consideration.

One obvious response to these problems is to sample words taking account of frequency of usage. Terman and Merrill (1937) arranged their sample of words in order of difficulty. When the subject failed at six consecutive words, the vocabulary test was stopped. Dupuy (1974) recommends a similar procedure. Time can be saved by such a procedure, but vocabulary size is likely to be underestimated, since most people would know a few rare words.

The characteristics of the two major, current word frequency compilations available (Carroll, Davies, & Richman, 1971; Kučera & Francis, 1967) suggest a problem with frequency sampling. These analyses indicate that the distribution of words is highly unbalanced, a conclusion reached over 25 years ago by Horn (1954), who calculated that about 2,000 word types will account for about 95% of "running words in adult writing;" 3,000 for 96.9%, 4,000 for 97.8%, and 10,000 for 99.4%. At the low frequency end of the scale, there is a tail that approaches infinity. Even in a huge corpus, a vast number of words appear only once, twice, or not at all. Of the 86,741 word types listed by Carroll, Davies, and Richman (1971) from a corpus of over 5 million running words, 35,079 (40.44%) appeared once. Kučera and Francis (1967) found 44.72% of the words appeared once in a sample of over one million words. So, if the test is short, the students run the risk of not being able to show that they

know several medium frequency words, since there will be such a large proportion of rare words in the sample. A resolution of this issue is important, since a frequency-based sampling technique seems the most accessible method for overcoming the problems of simple random sampling.

Frequency is a characteristic of a word which probably is very strongly related to the chances that the word will be known. One exception to this may be the case of compound words. While the root of the word may be very common and well-known, a certain affix-root compound may be very infrequent, but almost equally well-known if the affix is familiar. So one analytic approach to the relationship between word frequency and knowledge would entail the use of "family" frequency, that is, the frequency of the root word and all its compounds and derivatives. We might expect, then, that the relationship of this index of frequency of usage to the chances of the words being known would be more reliable.

Indeed, we are willing to go further and speculate that the relationship between family frequency and the chances of knowing a word resembles the curve presented in Figure 1. In terms of breadth of knowledge, we would expect a ceiling at the upper end of the frequency scale: most

---

Insert Figure 1 about here

---

people know all of the very common words. Other aspects of the curve would differentiate individuals: The point at which the curve dropped from the plateau level, and the slope of the function probably are the

two measures that would capture the important individual differences. Even for children, we might best think of the curve leveling out as the words become very infrequent, since it is likely that, from their hobbies, interests, or the occupation of their parents, most children would know some very rare words. Nevertheless we have drawn the lower portion of the curve as a broken line since we are less sure about the relationship in this area.

In summary, a good test of word knowledge would present the student with a large number of words, sampled liberally from the whole range of word frequency. Techniques should be developed which allow accurate estimation of the relationship of a given student's probability of knowing a word and the frequency of the word's family.

Determining That a Word is in a Person's Vocabulary

Four sorts of test formats have been employed in attempts to assess breadth of vocabulary knowledge: (a) multiple choice; (b) constructed answer, in which the student attempts to give a definition, a synonym, an illustration, or use of word in a sentence or phrase; (c) yes/no judgments, in which the student checks the words in a list that he or she knows; and (d) matching, where the student pairs off words with their synonyms.

The question that needs examination is which of these methods will be of most theoretical and practical value as a measure of vocabulary. We will discuss issues relating to three of these methods: multiple choice,

constructed answers, and yes/no judgments. Since the points raised about the multiple choice format apply even more cogently to matching, the latter will not be dealt with separately.

Multiple choice methods. People often possess partial knowledge of words. In these instances the items' distractors become crucial. An individual may select the correct synonym for platitude from the choices: (a) duck-billed mammal, (b) praise, (c) commonplace remark, (d) flatness. He may make the correct selection because he has heard the word used in reference to an utterance and with a negative connotation. This information, however, may not enable him to select correctly from (a) commonplace remark, (b) nonsense, (c) irrelevant question, (d) insult. The set of choices constrains the individual's response to different degrees, and different policies for generating distractors will, of course, lead to differences in performance.

Lepley (1955, 1965), for example, constructed two forms of a synonym test, one employing distractors from the same semantic category as the target and another which used distractors from semantically diverse categories. Lepley (1965) found equal reliability but significantly superior performance on the version requiring only gross discriminations. Lepley's results illustrate the influence of the distractors.

The multiple choice format is currently the most widely used in standardized vocabulary testing (e.g., Stanford Achievement Tests, 1973; Metropolitan Achievement Tests, 1970; California Achievement Tests, 1977).

The principal complaint raised here so far is that the distractors cannot avoid constraining the subject's response. If the purpose of the test is to provide data only on relative performance of people, not on absolute level of performance, then the distractors can be, and usually are, chosen to maximize the discriminating power of the item. If one is interested in vocabulary size, then this policy will not do.

Many vocabulary tests (e.g., Stanford, 1973) use sentence completion in a multiple choice format. The problems already mentioned apply even when the test simulates a real encounter with the target word. In addition, the question of the effects of various amounts of contextual support on estimated vocabulary size with groups of words that vary in frequency of usage has not been studied. There is research that suggests that individuals vary not only in the size of their reading vocabularies but also in their ability to use context to deduce the meanings of unknown and partly known words (Pearson & Studt, 1975; Mason, Knisely, & Kendall, 1978).

A tricky problem with the multiple choice format is that young children may not consider all the distractors (Asher, 1978; Brown, 1975; Vurpillot, 1968). They will often choose the first or second alternative if it makes reasonable enough sense. The test-taking strategies of older children on multiple choice tests are not yet well characterized, but there quite probably are strategic components of good performance which serve to increase spuriously the relationship between a multiple choice vocabulary test and other achievement or intelligence tests in the same

format. An insidious possibility is that some of the apparent growth in vocabulary knowledge over the elementary school years is really attributable to the acquisition of more sophisticated test-taking skills.

In conclusion, the multiple choice format is the most popular one. It makes relatively efficient use of examinee time and must be reasonably valid, otherwise the strong relationships between performance on such tests and other measures of linguistic competence, summarized at the beginning of this paper, would not have been obtained. The chief complication with the multiple choice format, when one wants absolute measures of vocabulary knowledge, is how to choose distractors. A further problem is that multiple choice tests may make demands on strategic knowledge in which young and poor readers are deficient.

Constructed answer measures. To overcome the problem of selecting distractors, several researchers, notably Seashore (1933), Smith (1941), and Terman and Merrill (1937), have used a constructed answer format, in which the subject reads or hears the target word and then writes or tells a definition of it, uses it in a sentence, gives a synonym for it, or in some other way provides an indication of its sense and reference. Subjects can be encouraged to do any one of these things just so long as the experimenter is convinced the word is "known." This format is capable of dealing with a variety of levels of knowing a word and avoids the issue of distractors. There are, however, two substantial problems with constructed answer measures: the problem of scoring the answers and the problem of what is really being measured.

In the written format, in particular, a constructed answer measure is confounded by factors such as spelling ability, sentence construction ability, and even the ability to write legibly, all of which may discourage a subject from elaborating on a word used or understood in conversation. A slightly more subtle problem, and one that is more difficult to control, resides in the fact that, if a liberal criterion is used and the subject is allowed a range of possible responses to a target word, then a particular strategy for responding may be adopted. The problem is that some words would be more easily explicated in a particular form. The word noun may be more easily explained through illustration than by definition, for instance. The research of Anglin (1970) and Wolman and Baker (1965) indicates that, up to the age of about 10-12 years, children tend to provide concrete definitions-by-illustration rather than by an inclusive term or synonym. It is entirely possible that, depending on scoring criteria, the preference at a different age for certain explanatory strategies could produce spurious estimates of the rate of vocabulary growth.

A really vexing problem is how liberally to score answers. How does one score synonyms in relation to apt illustrations or perfect usage in a sentence? In many instances, partial knowledge is displayed. In one of our own recent testing sessions, it became clear that many fifth grade students had partial knowledge of the word forbid. Several students knew that it had something to do with not being permitted to do something but did not have as part of their knowledge the fact that forbid is used

in imperative speech acts. We soon realized that, in this case, we needed to ask for its use in a sentence. We have found other more subtle and difficult cases of partial knowledge. For the word propelled, there was no problem in the students' recognition of the word because of their knowledge of propeller. When probed about the function of a propeller, many came close to generating the notion of propulsion on the theory that it would be strange to have a big round blade going around on the front of a plane unless it served some fairly fundamental purpose--and what planes do is move.

Some words have no near synonyms. There are other instances when the only synonym is a less frequent word than the target. In such cases, the subject is being asked to produce a rare word in order to show that a common word is known.

There are some almost irresistible tendencies displayed by an examiner when administering a test with a constructed answer format. After a few children have been tested, the examiner develops a sense of which words are easy and which are difficult. It requires conscious effort to avoid expecting more explanation of the difficult words and less for the easy words. If every subject has known chair and the current subject pats the seat of his stool as a response, then the tendency is to award full marks. If he pats the wall for edifice, however, he might not score so well. Similarly there is an urge to expect more elaborated responses from older subjects. The preschooler who tells you that an automobile "goes brrrrrmmm" will strike you more favorably than the

college sophomore who gives you the same answer. In addition, the experimenter will witness explanations of words which entail subtle non-verbal as well as verbal cues. Young children typically employ hand movements, facial expressions, and gestures in their communications especially when dealing with words that are a little difficult for them.

The horns of the dilemma are these: Stringent, operational, adult-like standards for evaluating whether a response indicates a word is known will confound what is supposed to be a measure of breadth of vocabulary knowledge with expository ability. Looser, more flexible standards will confound the measure with the subjective judgment of the examiners, which may change from word to word, subject to subject, and occasion to occasion.

So the liabilities of the constructed answer method are both logistical and substantial. It is inefficient per unit of testing and scoring time, and it seems to rely on often subtle intuitions on the part of the examiner, especially when the person displays partial knowledge of an item.

Yes/no format. The final format to be considered is the yes/no method. In this format, the student simply indicates whether or not the meaning of a word is known. Two of the major difficulties that have arisen consistently in the discussion of the other two major formats are the problem of what the test really measures and the need to present a large number of words chosen from a wide frequency range.

The checking format can satisfy the second criterion admirably, but problems of validity arise. Sims (1929) concluded:

> The writer is inclined to believe that a good guess as to whether or not a child knows the meaning of a word is almost as satisfactory a method of determining vocabulary as checking tests. The relative simplicity of such a measure, the ease of preparation and administration should not blind one to its invalidity. (p. 96)

Chall and Dale (1950) reported that the average tendency to overestimate word knowledge in the yes/no format over and above the definition format amounted to about 11%, and was more pronounced for rare words.

It ought to be no real surprise that a yes/no test uncorrected for guessing and risk-taking would give inflated estimates of vocabulary size and would correlate poorly with other measures. Consider the yes/no task from the point of view of the test taker. Some individuals may deny that they know the word gold because they do not know its atomic weight, while others will agree they know it because they have a feeling that it can be used to refer to a color. How do we take account of differences in the degree of confidence individuals must have before they are willing to say, "Yes, I know that word"?

One method is to mix in some nonsense words, which allows one to determine how often the student is guessing. The student who says "yes" to a lot of the nonsense words is acting without much reliable information. We are currently analyzing data collected from elementary and high school

subjects on large numbers of words. The students responded yes or no to a mixture of many English words and almost as many nonsense words. Later they completed standardized multiple choice questions on the real words. Our preliminary analyses have indicated that yes/no scores adjusted for guessing and risk-taking correlate highly with multiple choice performance. We later interviewed the subjects individually about a subset of the words. The data suggest that a value derived from the yes/no task gives a better estimate of true word knowledge than performance on the standardized multiple choice test.

The fact that words have multiple meanings poses a problem for the yes/no task, since presumably a person will check "yes" if he or she knows any meaning of a word. This is not a small problem. According to Lovell (1941), 43% of the words used by Seashore and Eckerson (1940) had multiple meanings. Recently, Balch (cited in Johnson & Pearson, 1978, p. 17) has reported that from 23% to 42% of the words in six widely used basic vocabulary lists have multiple meanings. In other recent research, Mason, Knisely, and Kendall (1978) have shown that children are much less likely to know the secondary than the primary meaning of words used in their secondary sense in a popular basal series. It is apparent that the yes/no format is not suitable for distinguishing which of the meanings of a word are known. When that is the goal, some other method of assessment is required.

In summary, the great attraction of the yes/no format is that it permits the presentation of a very large number of words in a given

interval of examinee time. Compared to the multiple choice format, it reduces somewhat the burden of preparing distractors and, compared to constructed answer formats, it sidesteps vagaries of scoring. The notable problem with the yes/no task is that scores of individuals will be influenced markedly by differences in tendency to take risks in the face of uncertainty. If this problem can be solved, the yes/no task might be very useful for assessment of breadth of word knowledge.

Conclusion

While current research demonstrates the importance of such factors as a reader's perspective on a text (Pichert & Anderson, 1977) and test structure (Meyer, 1975; Mandler & Johnson, 1977), it is also clear that word knowledge is a requisite for reading comprehension: People who do not know the meanings of very many words are most probably poor readers. There are serious gaps in our understanding of why this is true and of how word knowledge grows throughout the life span. Filling those gaps promises to be both an intellectual and a practical challenge of considerable importance. We judge that a critical first step is the development of improved methods of assessing breadth of vocabulary knowledge. It is only after some refinement has been achieved at this level that models of lexical development and instructional programs can be based on realistic expectations about the acquisition of word meanings.

We conclude our review of vocabulary knowledge and vocabulary size with the realization that, since the turn of the century, a tremendous

amount of energy has been put into answering the question, "How many words does an individual know?" We have come to wonder if this question is properly framed. The nature of language may make it unanswerable and thus, for scientific purposes, irrelevant. Empirical methods may be able to generate useful indices such as that discussed earlier--the relationship of the individual's knowledge of words to word frequency. To produce a single value from performance on a sample to represent total vocabulary size, however, may be an exercise that relies too heavily on the assumption of a static population of words and on an overly restrictive view of how we generate and use them.

References

Allen, L. R., & Garton, R. F. The influence of word knowledge on the word frequency effect in recognition memory. Psychonomic Science, 1968, 10, 401-402.

Ames, W. S. The understanding vocabulary of first grade pupils. Elementary English, 1964, 41, 64-68.

Anglin, J. M. The growth of word meaning. Cambridge, Mass.: M.I.T. Press, 1970.

Asher, S. R. Referential communication (Tech. Rep. No. 90). Urbana: University of Illinois, Center for the Study of Reading, June 1978.

Bayer, M. L. Primary grade understanding vocabulary as measured by an orally administered Basic Word Vocabulary Test. Unpublished master's thesis, Rutgers University, 1976. (ERIC Document Reproduction Service, No. ED 131 417)

Becker, W. C. Teaching reading and language to the disadvantaged--what we have learned from field research. Harvard Educational Review, 1977, 47, 518-543.

Bonser, F. G., Burch, L. H., and Turner, M. R. Vocabulary tests as measures of school efficiency. School and Society, 1915, 2, 714-718.

Bormuth, J. R. Readability: A new approach. Reading Research Quarterly, 1966, 1, 79-132.

Botzum, W. A. A factorial study of reasoning and closure factors. Psychometrika, 1951, 16, 361-386.

Brandenburg, G. C. Psychological aspects of language. Journal of Educational Psychology, 1918, 9, 313-332.

Brown, A. L. Recognition, reconstruction and recall of narrative sequences by preoperational children. Child Development, 1975, 46, 156-166.

California Achievement Tests. Monterey, Cal.: CTB/McGraw-Hill, 1977.

Carroll, J. B., Davies, P., & Richman, B. Word frequency book. Boston: Houghton Mifflin, 1971.

Chall, J. S., & Dale, E. Familiarity of selected health terms. Educational Research Bulletin, 1950, 39, 197-206.

Clark, E. V. What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), Cognitive development and the acquisition of language. New York: Academic Press, 1973.

Clark, H. H., & Clark, E. V. Psychology and language. New York: Harcourt Brace Jovanovich, 1977.

Clark, N. L. Hierarchial structure of comprehension skills. A.C.E.R., Hawthorn, Victoria, Australia (2 vols.), 1972.

Clifford, G. J. Words for school: The applications in education of the vocabulary researches of E. L. Thorndike. In P. Suppes (Ed.), Impact of research on education: Some case studies. Washington, D.C.: National Academy of Education, 1978.

Coleman, E. B. Developing a technology of written instruction: Some determiners of the complexity of prose. In E. Z. Rothkopt & P. E. Johnson (Eds.), Verbal learning research and the technology

of written instruction. New York: Columbia University, Teacher's College Press, 1971.

Cuff, N. B. Vocabulary tests. Journal of Educational Psychology, 1930, 21, 212-220.

Davis, F. B. Fundamental factors of comprehension in reading. Psychometrika, 1944, 9, 185-197.

Davis, F. B. Research in comprehension in reading. Reading Research Quarterly, 1968, 3, 499-545.

Dolch, E. How much word knowledge do children bring to grade 1? Elementary English Review, 1936, 13, 177-183.

Dupuy, H. P. The rationale, development, and standardization of a basic word vocabulary test. Washington, D.C.: U. S. Government Printing Office, 1974. (DHEW Publication No. HRA 74-1334)

Elwood, M. I. A preliminary note on the vocabulary test in the revised Stanford-Binet Scale. Journal of Educational Psychology, 1939, 30, 632-634.

Fillmore, C. J. The future of semantics. In R. Austerlitz (Ed.), The scope of American linguistics. Amherst, Mass.: Ridder Press, 1975.

Fleisher, L. S., & Jenkins, J. R. Effects of contextualized and decontextualized practice conditions on word recognition (Tech. Rep. No. 54). Urbana: University of Illinois, Center for the Study of Reading, July 1977.

Gentner, D. Evidence for the psychological reality of semantic components: The verbs of possession. In D. A. Norman, D. E. Rumelhart, & the LNR research group, Explorations in cognition. San Francisco: Freeman, 1975.

Gerlach, F. M. Vocabulary studies--false definition test. Colorado College Studies in Education and Psychology, No. 1, 1917.

Gillette, J. M. Extent of personal vocabularies and culture control. Science Monthly, 1927, 25, 451-467.

Goodman, K. S. Behind the eye: What happens in reading. In H. Singer & R. B. Ruddell (Eds.), Theoretical modes and processes of reading. Newark, Del.: IRA Publication, 1976.

Hall, W. S., & Tirre, W. C. The communicative environment of young children: Social class, ethnic and situational differences (Tech. Rep. No. 125). Urbana: University of Illinois, Center for the Study of Reading, May 1979.

Hartman, G. W. Further evidence on the unexpected large size of recognition vocabularies among college students. Journal of Educational Psychology, 1946, 37, 436-439.

Holley, C. E. Holley sentence vocabulary scale - grades 3 - 12. Bloomington, Ill.: Public School Publication Co., 1919.

Horn, E. Teaching spelling. Washington, D.C.: A.E.R.A., National Educational Association, 1954.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Hunt, E. Mechanics of verbal ability. Psychological Review, 1978, 85, 109-130.

Johnson, D. D., & Pearson, P. D. Teaching reading vocabulary. New York: Holt, Rinehart and Winston, 1978.

Kirkpatrick, E. A. Vocabulary test. Popular Science Monthly, 1907, 70, 157-164.

Klare, G. R. Assessing readability. Reading Research Quarterly, 1974-1975, 10, 62-102.

Kučera, H., & Francis, W. N. Computational analysis of present-day American English. Providence, R.I.: Brown University Press, 1967.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Labov, W. The boundaries of words and their meanings. In C. J. Bailey & R. Shuy (Eds.), New ways of analyzing variation in English. Washington, D.C.: Georgetown University Press, 1973.

Lepley, W. M. Rationale, construction, and preliminary try-out of the synonym vocabulary test. Journal of Psychology, 1955, 39, 215-225.

Lepley, W. M. The synonym vocabulary test: Form III, Journal of Psychology, 1965, 59, 109-112.

Lewinski, R. J. Vocabulary and mental measurement: A quantitative investigation and review of research. Journal of Genetic Psychology, 1948, 72, 247-281.

Lovell, G. D. Interrelations of vocabulary skills: Commonest versus multiple meanings. Journal of Educational Psychology, 1941, 21, 67-72.

Mahan, H. C., & Witmer, L. A note on the Stanford-Binet vocabulary test. Journal of Applied Psychology, 1936, 20, 258-263.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 1977, 9, 11-151.

Mason, J. M., Knisely, E., & Kendall, J. Effects of polysemous words on sentence comprehension (Tech. Rep. No. 85). Urbana: University of Illinois, Center for the Study of Reading, May 1978.

McNemar, Q. The revision of the Stanford-Binet scale. Boston: Houghton-Mifflin, 1942.

Metropolitan Achievement Tests. New York: Harcourt Brace Jovanovich, 1970.

Meyer, B. J. F. The organization of prose and its effects on memory. Amsterdan: North Holland, 1975.

Michael, W. B., Zimmerman, W. S., & Guilford, J. P. An investigation of the nature of the spatial relations and visualization factors in two high school samples. Education and Psychology Measurement, 1951, 11, 561-577.

Miller, G. A. The acquisition of word meaning. Child Development, 1978 49, 999-1004.

Miller, G. A., & Johnson-Laird, P. N. Language and perception. Cambridge, Mass.: Belknap Press, 1976.

Miner, J. B. Intelligence in the U.S. New York: Springer, 1957.

Pearson, P. D., & Studt, A. Effects of word frequency and contextual richness on children's word identification abilities. Journal of Educational Psychology, 1975, 67, 89-95.

Perfetti, C. A., & Lesgold, A. M. Coding and comprehension in skilled reading and implications for reading instruction. In L. B. Resnick & P. Weaver (Eds.), Theory and practice in early reading. Hillsdale, N.J.: Erlbaum, in press.

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. Journal of Educational Psychology, 1977, 69, 309-315.

Putnam, H. The meaning of 'meaning.' In H. Putnam (Ed.), Mind, language, and reality. Cambridge, Mass.: Cambridge University Press, 1975, 215-271.

Raven, J. C. The comparative assessment of intellectual ability. British Journal of Psychology, 1948, 39, 12-19.

Rosch, E. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), Cognitive development and the acquisition of languages. New York: Academic Press, 1973.

Rosch, E., & Mervis, C. B. Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 1975, 1, 573-605.

Seashore, R. H. Measurement and analysis of extent of vocabulary. Psychology Bulletin, 1933, 30, 709-710.

Seashore, R. H., & Eckerson, L. D. The measurement of individual differences in general English vocabularies. Journal of Educational Psychology, 1940, 31, 14-38.

Shibles, B. H. How many words does a first grade child know? Elementary English, 1959, 31, 42-47.

Silvestri, S., & Silvestri, R. Developmental analysis of the acquisition of compound words. Language, speech and hearing services in the school, 1977, 8, 217-221.

Sims, V. M. The realiability and validity of four types of vocabulary test. Journal of Educational Research, 1929, 20, 91-96.

Smith, M. E. An investigation of the development of the sentence and the extent of vocabulary in young children. University of Iowa studies in child welfare, 1926, 3, 92.

Smith, M. K. Measurement of the size of general English vocabulary through the elementary grades and high school. General Psychological Monographs, 1941, 24, 311-345.

Spache, G. The vocabulary tests of the revised Stanford-Binet as independent measures of intelligence. Journal of Educational Research, 1943, 36, 512-516.

Stanford Achievement Tests. New York: Harcourt Brace Jovanovich, 1973.

Terman, L. M. The measurement of intelligence. Boston: Houghton-Mifflin, 1916.

Terman, L. M. Vocabulary test as a measure of intelligence. Journal of Educational Psychology, 1918, 9, 452-466.

Terman, L. M., & Merrill, M. A. Measuring intelligence. Boston: Houghton-Mifflin, 1937.

Vurpillot, E. Judging visual similarity: The development of scanning strategies and their relation to differentiation. Journal of Experimental Child Psychology, 1968, 6, 632-650.

Wechsler, D. Manual Wechsler intelligence scale for children. New York: Psychological Corporation, 1949.

White, M. L. Mental age norms for vocabulary scores on the 1937 Stanford-Binet. Psychological Records, 1942, 5, 159-169.

Wittgenstein, L. Philosophical investigations. New York: Macmillan, 1953.

Wolman, R. N., & Baker, E. N. A developmental study of word definitions. Journal of General Psychology, 1965, 107, 159-166.

Wrigley, C., Saunders, D. R., & Newhaus, J. O. Application of the quatrimax method of rotation to Thurstone's P.M.A. study. Psychometrika, 1958, 23, 115-170.

Footnote

[1]We are indebted to Charles Fillmore for this example.

Table 1

Correlations of Various Vocabulary Tests With Tests of General Intelligence

Vocabulary Measure	Intelligence Measure	Subjects	N	r	Source
Terman, 1916	Binet (1916)	School children	631	.91	Terman (1918)
Terman, 1916	Binet (1916)	School children	269	.87	Mahan & Witmer (1936)
Terman, 1937	Binet (1916)	School children	65	.92	Spache (1943)
Terman, 1937	Binet (1916)	School children	1161	.98	Elwood (1939)
Terman, 1937	Binet (1916)	School children	753	.86	White (1942)
Terman, 1937	Binet (1916)	Standardization sample, ages 8-18	710	.71 to .86	McNemar (1942)
Wechsler	Wechsler	Adult males	1000	.82	Lewinski (1948)
Wechsler	WISC	Standardization sample, ages 7.5, 10.5, 13.5	600	.71 .87 .78	Wechsler (1949)
Raven	Binet	School children	150	.93	Raven (1948)
Dupuy	Various tests	School children	2397	.76	Dupuy (1974)
	Stanford Achievement Tests (1973) (vocabulary with total achievement test scores)	Standardization samples Grade 2 3 4 5 6 8	275,000 over grades and geog. locale	.82 .79 .80 .80 .83 .89	Stanford Achievement Test (1973)

Note. Adapted from Miner, (1957).

Table 2

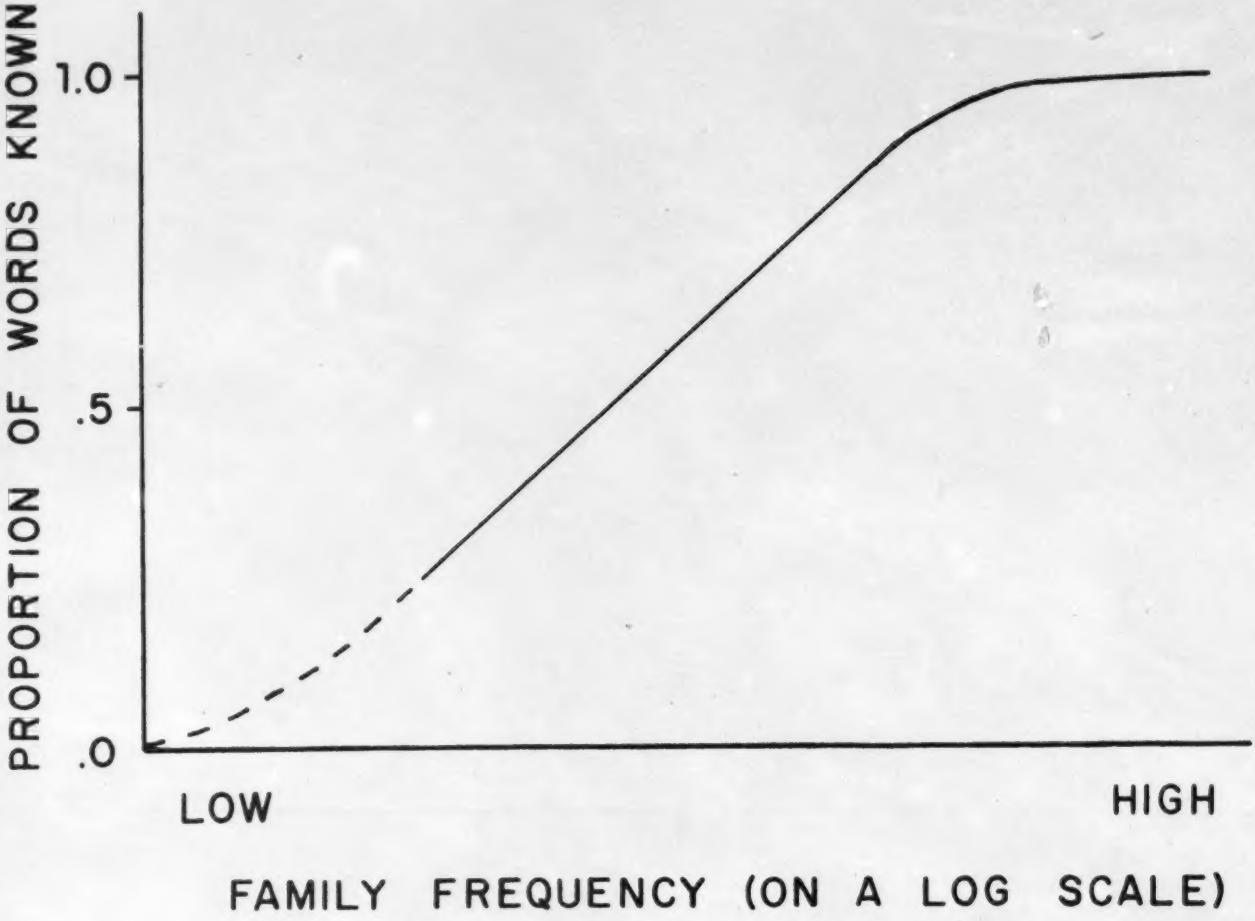
Some Previous Estimates of Total Vocabulary Size at Selected Grades

Grade	Source	Estimate
1st	M. E. Smith (1926)	2,562
	Dolch (1936)	2,703
	Ames (1964)	12,400
	M. K. Smith (1941)	17,000
	Shibles (1959)	26,000
3rd	Dupuy (1974)	2,000
	Holley (1919)	3,144
	Terman (1916)	3,600
	Brandenburg (1918)	5,429
	Kirkpatrick (1907)	6,620
	Cuff (1930)	7,425
	M. K. Smith (1941)	25,000
7th	Dupuy (1974)	4,760
	Terman (1916)	7,200
	Holley (1919)	8,478
	Kirkpatrick (1907)	10,666
	Brandenburg (1918)	11,445
	Cuff (1930)	14,910
	Bonser, et al. (1915)	26,520
	M. K. Smith (1941)	51,000
College sophomore	Seashore (1933)	15,000
	Kirkpatrick (1907)	19,000
	Seashore & Eckerson (1940)	60,000
	Gerlach (1917)	85,300
	Gillette (1927)	127,800
	Hartman (1946)	200,000

Note. Adapted from Seashore and Eckerson, 1940, and Bayer 1976.

Figure Caption

Figure 1. Possible relationship between likelihood word meanings are known and frequency of usage.


PROPORTION OF WORDS KNOWN
1.0
.5
.0
LOW
HIGH
FAMILY FREQUENCY (ON A LOG SCALE)

CENTER FOR THE STUDY OF READING

READING EDUCATION REPORTS

No. 1: Durkin, D. ***Comprehension Instruction—Where are You?,*** October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., PC-$1.82, MF-$.83)

No. 2: Asher, S. R. ***Sex Differences in Reading Achievement,*** October 1977. (ERIC Document Reproduction Service No. ED 145 567, 30p., PC-$3.32, MF-$.83)

No. 3: Adams, M. J., Anderson, R. C., & Durkin, D. ***Beginning Reading: Theory and Practice,*** November 1977. (ERIC Document Reproduction Service No. ED 151 722, 15p., PC-$1.82, MF-$.83)

No. 4: Jenkins, J. R., & Pany, D. ***Teaching Reading Comprehension in the Middle Grades,*** January 1978. (ERIC Document Reproduction Service No. ED 151 756, 36p., PC-$3.32, MF-$.83)

No. 5: Bruce, B. ***What Makes a Good Story?,*** June 1978. (ERIC Document Reproduction Service No. ED 158 222, 16p., PC-$1.82, MF-$.83)

No. 6: Anderson, T. H. ***Another Look at the Self-Questioning Study Technique,*** September 1978. (ERIC Document Reproduction Service No. ED 163 441, 19p., PC-$1.82, MF-$.83)

No. 7: Pearson, P. D., & Kamil, M. L. ***Basic Processes and Instructional Practices in Teaching Reading,*** December 1978. (ERIC Document Reproduction Service No. ED 165 118, 29p., PC-$3.32, MF-$.83)

No. 8: Collins, A., & Haviland, S. E. ***Children's Reading Problems,*** June 1979.

No. 9: Schallert, D. L., & Kleiman, G. M. ***Some Reasons Why Teachers are Easier to Understand than Textbooks,*** June 1979.

No. 10: Baker, L. ***Do I Understand or Do I not Understand: That is the Question,*** July 1979.

No. 11: Anderson, R. C., & Freebody, P. ***Vocabulary Knowledge and Reading,*** August 1979.